

# Evaluación de una Estrategia de Expansión Local Conservadora en Recuperación de Información Visual

## *Evaluating a Conservative Automatic Local Expansion Strategy for Visual Information Retrieval*

Sergio Navarro, Rafael Muñoz y Fernando Llopis

Departamento de Lenguajes y Sistemas Informáticos

Universidad de Alicante

Carretera Sant Vicent del Raspeig s/n

03690 Sant Vicent del Raspeig (Alicante)

[snavarro,rafael,llopis@dlsi.ua.es](mailto:snavarro,rafael,llopis@dlsi.ua.es)

**Resumen:** En este trabajo comparamos dos métodos de expansión de la pregunta en el área de la Recuperación de Información Visual (RIV): Probabilistic Relevance Feedback (PRF) y Local Context Analysis (LCA). La principal diferencia observada entre ambos métodos es que mientras PRF utiliza para la expansión las anotaciones correspondientes a las primeras imágenes de un ranking, LCA evita utilizar anotaciones correspondientes a imágenes no relevantes, situadas en esas primeras posiciones, mediante una heurística basada en coocurrencia. Los resultados muestran que LCA obtiene mejor precisión que PRF a medida que la precisión del ranking utilizado para la expansión es menor. Esta observación hace de LCA un método especialmente adecuado para su utilización con rankings de baja precisión como los devueltos por sistemas de RIV basados en el contenido de la imagen. Y así lo demuestran los buenos resultados obtenidos utilizando la variante multimodal de LCA, que es la única estrategia de expansión local que no daña a la diversidad de los resultados, y a su vez la que obtiene nuestros mejores resultados de precisión con el conjunto de consultas de la tarea ImageCLEFPhoto 2008 – 4º MAP y 5º P20 de las 1039 ejecuciones automáticas enviadas por los participantes –.

**Palabras clave:** Recuperación de Información Visual, Expansión Local, LCA, PRF

**Abstract:** In this paper we compare two query expansion methods in the Visual Information Retrieval (VIR) area: Probabilistic Relevance Feedback (PRF) and Local Context Analysis (LCA). The main difference observed between these methods is that while PRF assumes that annotations related to top-ranked images are relevant, LCA avoids to include terms from top-ranked non relevant images of the ranking using an heuristic based on cooccurrence. The experiment results show us that LCA increases its precision over PRF for those rankings with lowest precision. Thus, LCA demonstrates to be specially suitable for low precision rankings as the ones returned by the VIR systems based on the content of the image. Indeed, our multimodal LCA variation is the only one local expansion strategy which do not hurt the diversity of the results and the one which reach our best precision results with the ImageCLEFPhoto 2008 task query set – 4º MAP and 5º P20 within the 1039 automatics runs submitted by the participants –.

**Keywords:** Visual Information Retrieval, Relevance Feedback, LCA, PRF

## 1. Introducción

La gran variedad de formatos digitales existentes en la red y el boom de los contenidos multimedia, hacen necesario desarrollar y/o adaptar herramientas de búsqueda de información a las características de estos nuevos formatos como son el video y la imagen entre otros. Actualmente los buscadores comerciales de contenidos multimedia, como los

bien conocidos Youtube<sup>1</sup> o Flickr<sup>2</sup>, basan las búsquedas sólo en el texto que acompaña a la imagen o al video. El desarrollo de este tipo de herramientas está dentro del área de investigación de la Recuperación de Información Visual (RIV), que es donde se encuadra el trabajo que aquí presentamos. La RIV po-

<sup>1</sup><http://www.youtube.com>

<sup>2</sup><http://www.flickr.com>

dríamos decir que es un área específica dentro de la Recuperación de Información (RI), en la cual de hecho inicialmente se han utilizado sistemas de RI tradicionales sin ninguna adaptación específica a la RIV para llevar a cabo búsquedas utilizando las anotaciones que acompañan a las imágenes. Así, las colecciones que utilizan los sistemas de RIV están compuestas de colecciones de imágenes y de las anotaciones que describen a las mismas.

Históricamente en el área de la RIV se han utilizado dos enfoques para llevar a cabo la RI de imágenes: En los inicios de la RIV a finales de los años 70, los sistemas de RIV se basan en las anotaciones que acompañan a las imágenes, se trata pues de sistemas de RIV Basados en Texto (RIBT). Más tarde a principios de los años 90, en un intento de superar la dependencia de los sistemas de RIBT de la existencia de anotaciones textuales de una imagen para su indexación, surgen los sistemas de RIV Basados en el Contenido de la Imagen (RIBC) (Grubinger, 2007).

Finalmente en los últimos años a medida que las tecnologías utilizadas por los sistemas de RIBC han ido madurando, un tercer enfoque para afrontar el problema de la RIV ha surgido, se trata de los sistemas que combinan tecnologías basadas en texto y en imagen. En este contexto se han venido organizando competiciones como el ImageCLEF<sup>3</sup> que es una tarea específica de RIV que busca estimular el desarrollo de sistemas multimodales utilizando colecciones de imágenes con anotaciones reducidas de las mismas para la evaluación y comparación entre sistemas, y que se celebra en el marco de las competiciones anuales del CLEF<sup>4</sup>.

El objetivo de este trabajo es comparar dos estrategias textuales de expansión local de la consulta que han sido utilizadas, por un lado como técnicas de expansión local de la consulta en sistemas textuales de RIV y por el otro como estrategias de combinación multimodal en sistemas basados en texto e imagen. Las estrategias que comparamos son Probabilistic Relevance Feedback (PRF) y Local Context Analysis (LCA). PRF ha sido ampliamente utilizada como técnica de expansión local en sistemas textuales (Díaz-Galiano et al., 2007) y como estrategia de combinación multimodal (Gao et al., 2007) – utilizando las anotaciones asociadas al rank-

ing devuelto por un sistema de RIBC –. En cuanto a LCA, que es una estrategia de expansión local conservadora de la consulta que goza de menor popularidad que PRF, el único precedente que existe de utilización como técnica de combinación multimodal en RIV es el que presentamos en nuestra participación en la subtask Photo de la edición del 2008 del ImageCLEF. (Navarro, Llopis, y Muñoz, 2008).

A partir de los buenos resultados obtenidos con la versión multimodal de LCA en nuestra participación en la citada subtask Photo del ImageCLEF 2008, en este trabajo nos centramos en comparar PRF y LCA como métodos de expansión local en el ámbito de la RIBT. Nuestro objetivo es hallar evidencias en el mayor número posible de conjuntos de consultas y de colecciones de imágenes, que justifiquen los buenos resultados obtenidos por LCA en su versión multimodal.

El sistema que hemos utilizado para la RIBT es IR-n (Llopis, 2003), se trata de un sistema basado en pasajes, que ha demostrado en diferentes competiciones mejores resultados que los obtenidos por la mayoría de los sistemas de RIBT basados en documentos. (Navarro, Muñoz, y Llopis, 2008a; Navarro, Muñoz, y Llopis, 2008b).

El artículo está estructurado como sigue: en primer lugar presentamos las principales características del sistema IR-n centrándonos en los métodos de expansión local automática utilizados. A continuación describimos las colecciones utilizadas, los experimentos y los resultados obtenidos. Finalmente informamos de las conclusiones y de las propuestas de trabajos futuros extraídas a partir de los resultados.

## 2. Sistema IR-n

Para llevar a cabo los experimentos hemos utilizado IR-n, un sistema de recuperación de información basado en pasajes. Este tipo de sistemas tratan cada documento como un conjunto de pasajes, donde cada pasaje delimita una porción de texto del documento. Al contrario que los sistemas basados en documentos, los sistemas basados en pasajes permiten dar mayor relevancia a aquellos documentos donde los términos de la consulta aparecen en posiciones más cercanas entre sí (Llopis, 2003).

IR-n utiliza listas de palabras de parada

<sup>3</sup><http://www.imageclef.org>

<sup>4</sup><http://www.clef-campaign.org>

y stemmers para determinar que información de un documento es utilizada para la recuperación de información. Por un lado, la lista de palabras de parada de cada idioma contiene términos que son ignorados por el sistema por considerar que su presencia en la consulta o en un documento no es lo suficientemente significativa como para determinar si un documento es relevante para una consulta. Por otro lado, el stemmer de cada idioma es utilizado para obtener la raíz de cada término – eliminando sufijos y/o prefijos – tanto de los términos de la consulta en la fase de búsqueda como de los términos de los documentos en la fase de indexación de la colección. Una lista de los stemmers y las listas de palabras de parada utilizadas por IR-n puede ser consultada en [www.unine.ch/info/clef](http://www.unine.ch/info/clef).

Además IR-n permite seleccionar entre diferentes modelos de recuperación. El modelo de recuperación utilizado determina como se va a medir la similitud entre un texto y una consulta, para ello cada modelo utiliza una fórmula que valora el número de términos de la consulta que están presentes en el documento y el poder discriminador de cada uno de estos términos en la colección.

## 2.1. Expansión Local de la Consulta

En RIBT la estrategia de expansión local de la consulta más extendida es la de Probabilistic Relevance Feedback (PRF) (Robertson y Sparck Jones, 1977). Esta estrategia considera como relevantes los  $m$  primeros documentos devueltos en un ranking, a partir de los cuales extrae los  $j$  primeros términos mejor valorados según la Fórmula (1) aplicada a cada término  $t$ :

$$w_t = \frac{(m_t + 0,5) \cdot (n - n_t - m + m_t + 0,5)}{(m - m_t + 0,5) \cdot (n_t - m_t + 0,5)} \quad (1)$$

donde  $n$  es el número de documentos en la colección,  $n_t$  es el número de documentos en los que aparece el término  $t$ , y  $m_t$  es el número de documentos considerados relevantes en los que aparece  $t$ . De manera que  $w_t$  devuelve un valor mayor para aquellos términos cuya frecuencia entre los  $m$  primeros documentos del ranking es mayor que en toda la colección.

Aunque esta técnica puede empeorar los resultados en el caso de que la mayoría de

los documentos situados en las primeras posiciones del ranking no sean relevantes, los resultados alcanzados por los sistemas que la han utilizado en las conferencias TREC y CLEF muestran que en términos generales es una técnica efectiva (Xu y Croft, 2000), siendo de hecho la técnica de expansión local más utilizada por sistemas de RIV en anteriores ediciones del ImageCLEF (Gao et al., 2007) (Díaz-Galiano et al., 2007).

Una aproximación alternativa más conservadora que no parte de la asunción de que todos los documentos devueltos en las primeras posiciones del ranking son relevantes, es Local Context Analysis (LCA) (Xu y Croft, 2000). Esta estrategia al igual que PRF se basa en la frecuencia de aparición de los términos presentes en los  $n$  primeros documentos del ranking para la selección de términos a utilizar en la expansión, pero al contrario que la primera, trata de evitar la utilización de términos pertenecientes a documentos no relevantes, para ello da una valoración más alta a aquellos términos pertenecientes a documentos con mayor coocurrencia de términos con la consulta.

Así, los autores de LCA denominan conceptos a los términos de la expansión. Y utilizan una función  $f(c, Q)$ , Fórmula (2), para valorar cómo de adecuado es un concepto  $c$  para expandir una consulta  $Q$  basándose en las coocurrencias de  $c$  con los términos de la consulta en los  $n$  primeros documentos del ranking. Su Fórmula  $f(c, Q)$  viene dada por:

$$\begin{aligned} f(c, Q) &= \prod_{w_i \in Q} (\delta + co\_de(c, w_i))^{idf(w_i)} \\ co\_de(c, w_i) &= \frac{\log_{10}(co(c, w_i) + 1)idf(c)}{\log_{10}(n)} \\ co(c, w_i) &= \sum_{d \in S} tf(c, d)tf(w_i, d) \\ idf(c) &= \min(1, 0, \log_{10}(N/N_c)/5, 0) \end{aligned}$$

Donde  $w_1, w_2, \dots, w_m$  son los términos de la consulta  $Q$ ,  $N$  es el número de documentos en la colección,  $N_c$  el número de documentos que contienen el concepto  $c$ , y  $tf(c, d)$  y  $tf(w_i, d)$  son las frecuencias de  $c$  y  $w_i$  en un documento  $d$  respectivamente. Y finalmente  $\delta$  es utilizado simplemente como técnica de suavizado para evitar valores nulos – los autores recomiendan utilizar un valor de 0.4 –.

La principal limitación de este método basado en coocurrencia es que no siempre hay

coocurrencia entre un documento relevante y la consulta (Liu y Junzhong Gu, 2007). Este problema podría acusarse más en colecciones de imágenes, donde las anotaciones son reducidas, y por tanto la colección tiene reflejadas en sus documentos menos relaciones entre términos.

Los experimentos realizados por (Xu y Croft, 2000) demuestran que el rendimiento de LCA es más independiente que el de PRF ante diferencias en el número de documentos utilizados para la expansión. Lo que prueba su efectividad a la hora de desechar documentos no relevantes en las primeras posiciones del ranking. Contrariamente a lo que se podría esperar, esta estrategia de expansión local no goza de tanta popularidad como PRF.

Desde el punto de vista de la utilización de la expansión local como técnica de combinación multimodal, los trabajos previos del estado de la cuestión basados en expansión local textual como técnica de combinación multimodal han utilizado PRF, no habiendo precedentes de utilización de LCA, salvo nuestra participación en la tarea Photo del ImageCLEF 2008

Para implementar la estrategia de realimentación multimodal, nuestro sistema utiliza las  $n$  primeras anotaciones de las imágenes en el ranking devuelto por un sistema de RIBC externo y las  $i$  primeras anotaciones en el ranking devuelto por un sistema de RIBT obtenido con IR- $n$ . Para a continuación, extraer los  $t$  términos mejor valorados por el algoritmo utilizado – PRF o LCA –.

Indicar finalmente que IR- $n$  permite configurar si se utilizan los documentos o los pasajes más relevantes para la selección de términos de la expansión local de la consulta.

### 3. Experimentos

Para nuestros experimentos hemos escogido las colecciones de imágenes y los conjuntos de preguntas utilizados en las tres ediciones anteriores a la edición de este año del ImageCLEF. A continuación se comentan las características de cada una de las colecciones utilizadas.

- **St Andrews (Photo 2004)** (Reid, 1999). Se trata de una colección de fotografías históricas en blanco y negro, las anotaciones que acompañan a las imágenes se puede considerar que son de

alta calidad, sus autores son expertos en el área. Las mismas contienen un total de 8 campos en texto plano, que son: un título corto y otro largo, localización, descripción, fecha, fotógrafo, notas y categorías. Siendo los dos últimos especialmente ricos en información del contexto de la fotografía – información que un humano no podría extraer sólo observando la misma –. Esta colección es la que posee anotaciones de mayor calidad y extensión de las utilizadas.

- **IAPR TC-12 (Photo 2006 y Photo 2007 y 2008)** (Clough et al., 2006; Grubinger et al., 2007). Se trata de una colección de imágenes de lugares del mundo en su mayoría tomadas por una compañía de viajes. Las mayores diferencias con la de St. Andrews son que IAPR TC-12 utiliza imágenes de color y anotaciones en formato XML con los siguientes campos: título, descripción, notas, lugar y fecha.

Para la tarea del 2006 se proporciona una versión de la colección IAPR TC-12 a la que se le aplica un preproceso para reducir la calidad de las anotaciones, quedando como sigue: un 70 % de las imágenes poseen todos los campos completos, un 10 % no contienen descripción, otro 10 % no tienen ni descripción ni título, y finalmente otro 10 % no tiene anotaciones.

Para la tarea del 2007 se proporciona una versión para la cual las imágenes contienen todos los campos salvo el de descripción, que por otra parte es el más rico en semántica. Esto reduce considerablemente la cantidad de información textual de las anotaciones de la colección si lo comparamos con las anotaciones de la versión utilizada en el 2006.

La Tabla 3 muestra los datos más relevantes de cada colección, donde:

- **Colección:** Nombre de la colección junto a la tarea y las ediciones en que se utilizó como colección de test.
- **Idioma:** Idioma de la colección utilizada en nuestros experimentos.
- **N. Docs.:** Número de imágenes de la colección.

Colección	Idioma	N Docs	Media Pals	Media Fra
St Andrews (Photo 2004,2005)	Inglés	28.133	60.7	4.18
IAPR TC-12 (Photo 2006)	Inglés	20.000	27.46	2.32
IAPR TC-12 (Photo 2007,2008)	Inglés	20.000	12.93	2,6

Cuadro 1: Colecciones de Datos para RIV

- **Media Pals.:** Número medio de palabras que contienen las anotaciones de una imagen.
- **Media Fra.:** Número medio de frases que contienen las anotaciones de una imagen.

En cuanto a los conjuntos de consultas utilizadas, indicar que dependiendo de la edición de la competición, las consultas contienen un campo de título y un campo de narrativa o sólo un campo de título. En nuestros experimentos sólo hemos utilizado el campo de título de la pregunta, por considerar que no es realista utilizar la narrativa ya que un usuario suele utilizar consultas de reducido tamaño cuando utiliza un sistema de RIV. Por otro lado comentar que mientras que el conjunto de consultas de la edición del 2004 – 25 consultas – y de la edición del 2005 – 28 consultas – es diferente, para la edición del 2006 y del 2007 se utiliza el mismo conjunto de consultas – 60 consultas –, variando sin embargo las características de la colección utilizada.

A continuación se muestra la notación utilizada en las tablas de resultados para referirnos a cada uno de los parámetros del sistema:

- **Expansión Local (FB):** Indica el tipo de expansión local que se utiliza. Permite las versión textual de PRF y LCA y su versión multimodal PRFMM y LCAMM.
- **Parámetros para la Expansión Local:** Si  $E$  tiene valor 1, esto denota que se está utilizando expansión local basada en pasajes. Pero, si  $E$  tiene valor 2, la expansión local está basada en documentos. Además,  $N$  denota el número de documentos que la expansión local usará del ranking textual obtenido,  $N_{cbir}$  en el caso de estar utilizando expansión

multimodal denota el número de documentos que utilizará del ranking visual obtenido por un sistema de RIBC, y finalmente,  $T$  indica el número de términos a añadir a la consulta textual.

Para los experimentos hemos utilizado como modelo de recuperación Divergence From Randomness (DFR) (Amati y Van Rijsbergen, 2002), por ser el modelo de recuperación que mejores resultados obtuvo en nuestra participación para la colección en inglés en la tarea Photo del ImageCLEF 2007 (Navarro et al., 2008). Además, el tamaño de pasaje y los parámetros de DFR utilizados corresponden con los que mejores resultados han devuelto para cada colección sin la utilización de expansión local.

La estrategia seguida en los experimentos ha sido la de utilizar para los parámetros de configuración de la expansión local un rango de valores amplio que permita determinar que configuraciones son más apropiadas para cada técnica en relación al conjunto de preguntas y colección utilizada. Los valores utilizados para estos parámetros han sido de 5 a 100 documentos para el número de documentos utilizados para la consulta, y de 5 a 100 términos para el número de términos seleccionados para la expansión, además se han realizado experimentos utilizando tanto pasajes como documentos para la expansión.

Las siguientes tablas de resultados muestran para cada conjunto de experimentos el resultado del caso base – una ejecución sin expansión de la consulta – y las nueve mejores ejecuciones en orden creciente de MAP, utilizando el mejor MAP obtenido por PRF y LCA para esa misma configuración.

La Tabla 2 y la Tabla 3 muestran los resultados obtenidos en los experimentos realizados con la colección St. Andrews, podemos observar como para la misma colección con

E	N	T	PRF	LCA
			0.7262	0.7262
2	75	5	0.7087	<b>0.7345</b>
2	50	5	0.7254	<b>0.7355</b>
2	45	5	<b>0.7361</b>	0.7321
2	60	5	0.7129	<b>0.7368</b>
2	65	5	0.7180	<b>0.7375</b>
1	45	10	<b>0.7385</b>	0.7104
1	40	10	<b>0.7387</b>	0.7047
2	40	5	0.7395	<b>0.7404</b>
<b>2</b>	<b>40</b>	<b>10</b>	<b>0.7436</b>	0.7146

Cuadro 2: St. Andrews, 2004

E	N	T	PRF	LCA
			0.3493	0.3493
1	55	5	<b>0.3504</b>	0.3501
2	25	5	<b>0.3529</b>	0.3478
2	5	15	<b>0.3545</b>	0.3436
1	25	5	<b>0.3559</b>	0.3499
2	70	5	0.3246	<b>0.3566</b>
2	80	5	0.3206	<b>0.3580</b>
2	5	5	0.3455	<b>0.3596</b>
<b>2</b>	<b>10</b>	<b>5</b>	0.3533	<b>0.3607</b>
<b>2</b>	<b>50</b>	<b>5</b>	0.3307	<b>0.3607</b>

Cuadro 3: St. Andrews, 2005

diferentes conjuntos de consultas se obtienen resultados de precisión muy diferentes para los casos base utilizados, así las consultas del 2004 obtienen valores de precisión muy superiores a los alcanzados con las consultas del 2005.

Esto explica que para el conjunto de experimentos con un caso base con mayor precisión – consultas del 2004 – los resultados sean ligeramente superiores utilizando PRF,

E	N	T	PRF	LCA
			0.1988	0.1988
1	20	10	0.2126	<b>0.2140</b>
2	20	10	0.2146	<b>0.2191</b>
1	10	10	0.2174	<b>0.2218</b>
2	10	10	0.2186	<b>0.2222</b>
2	40	10	0.2025	<b>0.2226</b>
1	5	5	0.2119	<b>0.2233</b>
2	25	5	0.2004	<b>0.2241</b>
<b>2</b>	<b>5</b>	<b>5</b>	0.2126	<b>0.2242</b>

Cuadro 4: IAPR TC-12, 2006

E	N	T	PRF	LCA
			0.1544	0.1544
2	5	5	0.1808	<b>0.1810</b>
1	10	5	<b>0.1827</b>	0.1806
2	10	5	<b>0.1827</b>	0.1817
2	25	10	0.1811	<b>0.1834</b>
1	5	5	<b>0.1849</b>	0.1781
1	50	10	0.1675	<b>0.1884</b>
2	90	5	0.1452	<b>0.1898</b>
2	50	5	0.1551	<b>0.1911</b>
2	85	5	0.1444	<b>0.1918</b>
<b>2</b>	<b>45</b>	<b>5</b>	0.1598	<b>0.1923</b>

Cuadro 5: IAPR TC-12 No-Descriptions, 2007

y para los experimentos con precisión más baja – consultas del 2005 – los resultados sean considerablemente mejores utilizando LCA.

También podemos observar que mientras que PRF obtiene peores resultados cuando utiliza un número alto de documentos para llevar a cabo la expansión de la consulta, por el contrario LCA muestra resultados más independientes del número de documentos seleccionados para la expansión.

La Tabla 4 y la Tabla 5 muestran los resultados alcanzados para la colección del 2006 y 2007, podemos ver como afecta la reducción en el tamaño de las anotaciones a la precisión en estos experimentos. Además vemos como el hecho de que el caso base obtenga precisiones más bajas influye en que LCA amplíe la mejora de sus resultados respecto a PRF. Esto se explica porque en un ranking de baja precisión el número de imágenes no relevantes utilizadas para la expansión es mayor, lo cual daña más a los resultados obtenidos por PRF.

Se observa como LCA obtiene mejores resultados que PRF para todas las colecciones y conjuntos de preguntas, salvo para el conjunto de preguntas del 2004. Siendo en este último caso la diferencia de precisión entre PRF y LCA muy poco significativa. Por otro lado observamos como la diferencia obtenida entre LCA y PRF para el resto de colecciones se amplía a medida que la precisión del ranking devuelto por la ejecución del caso base obtiene una menor precisión.

Además, los resultados muestran que en general el número de documentos que LCA es capaz de manejar para la expansión es muy superior al número utilizado por PRF, con-

statándose así las observaciones apuntadas por los autores de LCA.

Finalmente, comparamos los resultados de estos experimentos con los alcanzados por las ejecuciones que utilizaron expansión local en nuestra participación en la tarea Photo del ImageCLEF 2008 (Navarro, Llopis, y Muñoz, 2008) – Tabla 6 –, en la cual se utilizó la colección IAPR TC-12 No-Descriptions y un subconjunto de las consultas utilizadas en la edición del 2007.

Podemos ver por un lado que a pesar de haber utilizado la misma colección y un subconjunto de las consultas utilizadas en los experimentos del 2007, la precisión alcanzada por el caso base en la participación del 2008 es superior a la obtenida por el caso base de los experimentos del 2007. Esto explica que en nuestra participación del 2008, PRF en su versión textual haya mejorado ligeramente los resultados obtenidos por LCA.

Por otro lado vemos que el mejor resultado multimodal ha sido alcanzado con LCA utilizando el ranking devuelto por un sistema de RIBC – este tipo de sistemas se caracterizan por devolver rankings de baja precisión –. Recaltar que en nuestra participación en la tarea del 2008 esta variante multimodal de LCA fue la única estrategia de expansión local que no dañó a la diversidad de los resultados y a su vez nuestra mejor ejecución en términos de precisión, obteniendo el 4º mejor MAP y el 5º mejor P20 de las 1039 ejecuciones automáticas enviadas por los participantes, a pesar de ser el único grupo que no utilizó la narrativa de la pregunta. Es importante tener en cuenta este dato, pues los resultados del único participante que envió ejecuciones con y sin narrativa demuestran que simplemente incluyendo la primera frase de la narrativa su sistema dobla los resultados de precisión e incrementa notablemente sus valores de diversidad (Demerdash, Kosseim, y Bergler, 2008).

Se observa también como mientras la mejor ejecución de PRF multimodal sólo es capaz de utilizar los 5 primeros documentos del ranking textual y los 5 mejores documentos del ranking visual, LCA en cambio en su mejor ejecución hace uso sólo de las imágenes devueltas en el ranking visual, siendo capaz además de manejar un número alto de imágenes para la expansión – 20 imágenes –.

FB	E	N	Ncbir	T	MAP
-					0.2362
LCA	1	5	0	5	0.2656
PRF	2	5	0	5	0.2699
PRFMM	2	5	5	5	0.2856
LCAMM	2	0	20	5	0.3436

Cuadro 6: ImageCLEFPhoto 2008

Revisando todos los experimentos llevados a cabo podemos afirmar que en todos ellos cuando el ranking utilizado para la expansión alcanza valores de precisión baja, LCA obtiene mejoras significativas respecto a PRF, mientras que cuando se utilizan rankings de precisión más alta PRF obtiene resultados ligeramente superiores a los de LCA.

#### 4. Conclusiones y Trabajos Futuros

El estudio de los experimentos realizados ha puesto de manifiesto una relación clara entre la precisión del ranking utilizado para la expansión y los resultados obtenidos por las estrategias de expansión evaluadas.

Tener constancia de esta relación, nos permite afirmar que LCA es una estrategia robusta que encaja perfectamente con los rankings de baja precisión con que se trabaja en RIV, especialmente los devueltos por sistemas de RIBC. Esta observación se refuerza por el hecho de que LCA es capaz de aprovechar la presencia de términos coocurrentes con la consulta en estos rankings, lo cual es una pista especialmente significativa de relevancia, pues contienen anotaciones de imágenes obtenidas sólo utilizando técnicas visuales.

Por ello creemos que de cara a trabajos futuros su utilización como estrategia de combinación multimodal es prometedora pues tal como algunos autores afirman el resto de estrategias de expansión local mejoran la precisión a costa de empeorar la diversidad de los resultados devueltos (Sanderson, 2008), mientras que LCA multimodal a demostrado no sólo ser la única que no perjudica a la diversidad sino que además ha obtenido nuestros mejores resultados de precisión. Esto no hace sino que demostrar que esta técnica es capaz de utilizar ambas fuentes de información – imagen y texto – de forma óptima, haciendo ascender en el ranking textual final sólo a las imágenes con anotaciones simi-

lares a las imágenes relevantes situadas en las primeras posiciones del ranking visual, en lugar de a las imágenes similares a las ya situadas en las primeras posiciones del ranking textual inicial.

## 5. Agradecimientos

Esta investigación ha sido llevada a cabo utilizando fondos procedentes del Gobierno español a través del proyecto TEXT-MESS (TIN-2006-15265-C06-01) y de la Unión Europea(UE) a través del proyecto QALL-ME (FP6-IST-033860).

## Bibliografía

- Amati, G. y C. J. Van Rijsbergen. 2002. Probabilistic Models of information retrieval based on measuring the divergence from randomness. *ACM TOIS*, 20(4):357–389.
- Clough, Paul, Michael Grubinger, Thomas Deselaers, Allan Hanbury, y Henning Müller. 2006. Overview of the imageclef 2006 photographic retrieval and object annotation tasks. En *Working Notes of the 2006 CLEF Workshop*, Alicante, Spain, Septiembre.
- Demerdash, Osama El, Leila Kosseim, y Sabine Bergler. 2008. CLaC at ImageCLEFPhoto 2008. En *on-line Working Notes, CLEF 2008*.
- Díaz-Galiano, M.C., M.A. García-Cumbreras, M.T. Martín-Valdivia, A. Montejo-Raez, y L.A. Ureña-López. 2007. Sinai at imageclef 2007. En *Working Notes of the 2007 CLEF Workshop*, Budapest, Hungary, Septiembre.
- Gao, Sheng, Jean-Pierre Chevallet, Thi Hoang Diem Le, Trong Ton Pham, y Joo Hwee Lim. 2007. Ipal at imageclef 2007 mixing features, models and knowledge. En *Working Notes of the 2007 CLEF Workshop*, Budapest, Hungary, Septiembre.
- Grubinger, Michael. 2007. *Analysis and Evaluation of Visual Information Systems Performance*. Ph.D. tesis, Engineering and Science Victoria University.
- Grubinger, Michael, Paul Clough, Allan Hanbury, y Henning Müller. 2007. Overview of the ImageCLEFphoto 2007 photographic retrieval task. En *Working Notes of the 2007 CLEF Workshop*, Budapest, Hungary, Septiembre.
- Liu, Haixue y Zhao Lv Junzhong Gu. 2007. Improving the Effectiveness of Local Context Analysis Based on Semantic Similarity. En *2007 International Conference on Convergence Information Technology (IC-CIT 2007)*.
- Llopis, Fernando. 2003. *IR-n: Un Sistema de Recuperación de Información Basado en Pasajes*. Ph.D. tesis, University of Alicante.
- Navarro, Sergio, Fernando Llopis, Rafael Muñoz Guillena, y Elisa Noguera. 2008. Analysing an approach to information retrieval of visual descriptions with ir-n, a system based on passages. *Advances in Multilingual and Multimodal Information Retrieval: 8th Workshop of the Cross-Language Evaluation Forum, CLEF 2007, Budapest, Hungary, September 19-21, 2007, Revised Selected Papers*, páginas 522–529.
- Navarro, Sergio, Fernando Llopis, y Rafael Muñoz. 2008. Different Multimodal Approaches using IR-n in ImageCLEFphoto 2008. En *on-line Working Notes, CLEF 2008*.
- Navarro, Sergio, Rafael Muñoz, y Fernando Llopis. 2008a. A Multimodal Approach to the Medical Retrieval Task using IR-n. En *on-line Working Notes, CLEF 2008*.
- Navarro, Sergio, Rafael Muñoz, y Fernando Llopis. 2008b. A Textual Approach based on Passages Using IR-n in WikipediaMM Task 2008. En *on-line Working Notes, CLEF 2008*.
- Reid, N. 1999. The photographic collections in st andrews university library. *Scottish Archives*, 5:83–90.
- Robertson, S. E. y K. Sparck Jones. 1977. Relevance weighting of search terms. *Journal of the American Society for Information Science*, 27(3):129–146.
- Sanderson, Mark. 2008. Ambiguous queries: Test collections need more sense. En *SI-GIR'08*, páginas 20–24, July.
- Xu, Jinxi y W. Bruce Croft. 2000. Improving the effectiveness of information retrieval with local context analysis. *ACM Trans. Inf. Syst.*, 18(1):79–112.